NETWORKERS

IPS-430
2891_05_2001_c1 1

CISCO SYSTEMS

Advanced Concepts and Developments in Quality of Service

Session IPS-430

Agenda

- **Introduction**
- **Differentiated Services (Diff-Serv)**
- **RSVP and Integrated Services (Int-Serv)**
- **The QoS Role of MPLS**
- **Conclusions**

Is QoS Necessary?

Cisco.com

- It's hard to argue against the "just throw bandwidth at it" argument
  - Low utilization is an effective tool for QoS
  - Providing low utilization everywhere and for all traffic regardless of need may be impractical or too costly, esp. on WAN links
  - "Abundant" BW has a habit of getting eaten, e.g. TCP apps use all the BW and voice is hosed
- We'll assume that BW isn't abundant everywhere and thus QoS is needed

IPS-430
2891_05_2001_c1 5

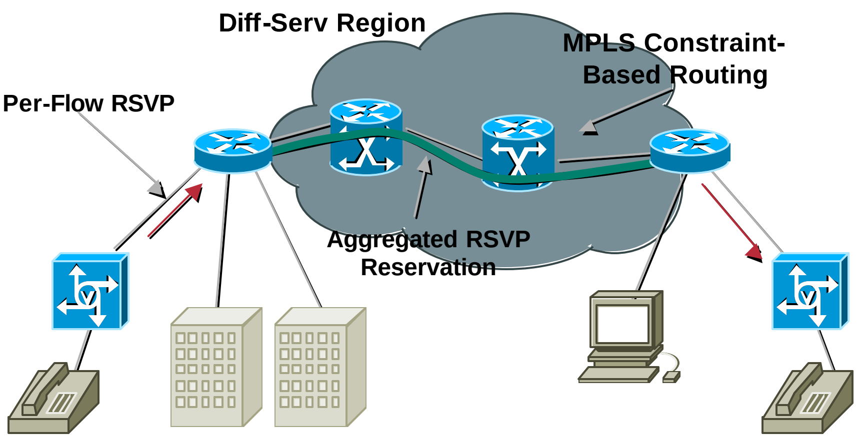
Why Those Topics?

Cisco.com

- Diff-Serv is the preferred QoS technology for large-scale deployment, e.g. ISPs
- Int-Serv and RSVP are not dead
  - In use to provide stronger guarantees and topology-aware admission control
  - Very useful for voice
- MPLS is being deployed widely by service providers
  - Its role in QoS is much debated
  - Often deployed in conjunction with Diff-Serv

IPS-430
2891_05_2001_c1 6

QoS Technologies



Objectives

- **Help you pick and choose among the various QoS technology offerings**
- **Present the trends and upcoming developments in IP QoS**
- **Help you impress friends with bleeding-edge QoS knowledge**

Why the IP focus?

- **IP is the dominant (inter)network layer protocol**
- **TCP is the dominant data transport**

  90–95% of Internet traffic uses TCP
- **VOIP is a growing market**
- **Heterogeneous link layers**

Differentiated Services (Diff-Serv)

Recent Developments

Diff-Serv Agenda

Cisco.com

- **Diff-Serv Recap**
- **The New “EF” (Expedited Forwarding)**
- **Worst-Case Bounds and Guarantees**
- **Services and Per-Domain Behaviors (PDBs)**

IPS-430
2891_05_2001_c1 11

Diff-Serv Overview

Cisco.com

- **Clearly the preferred QoS technology for ISPs today**
- **Near-minimal complexity**

  E.g. Can deploy DS with just 1 header bit and 2 “per-hop behaviors” (PHBs)
- **Edge behavior (classification, marking, policing etc.) + core behavior (PHBs) provides services**

  Variety of services from a single PHB

IPS-430
2891_05_2001_c1 12

Diff-Serv Contributions

- **Standardized definition of the "Diff-Serv Code Point" (DSCP)**

  (After years of confusion about TOS)

- **Small set of standard PHBs**

  EF, AF, etc.

- **An overall architecture for DS**

  Mostly formalizing ideas already in use

IPS-430
2891_05_2001_c1 13

Diff-Serv and Cisco

- **A subset of Diff-Serv functionality has been available in Cisco IOS for some time**
- **CAR, WRED, CBWFQ, LLQ are all Diff-Serv building blocks**
- **IOS originally used IP Precedence (3 bits)**

  Full 6-bit DSCP support now becoming available

IPS-430
2891_05_2001_c1 14

Defined PHBs

Cisco.com

- **Expedited Forwarding (EF)**
  - Dedicated low delay queue
- **Assured Forwarding (AF)**
  - n queues ? m drop preferences
- **Class Selector (CS)**
  - Backward compatible w/precedence
- **Default (best effort)**

IPS-430
2891_05_2001_c1 15

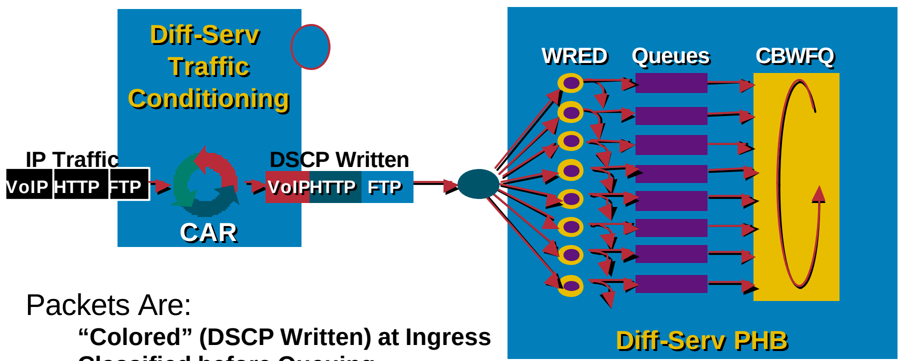
Example Service with Diff-Serv

Cisco.com

- **Classifier + token bucket policer at network edge**
  - Recognize and meter traffic in need of isolation; set DSCP = x
- **Dedicate a queue (and some bandwidth) to DSCP = x**
- **Effect is to run this traffic on its own logical network (with controlled utilization)**

IPS-430
2891_05_2001_c1 16

Putting It All Together



Packets Are:

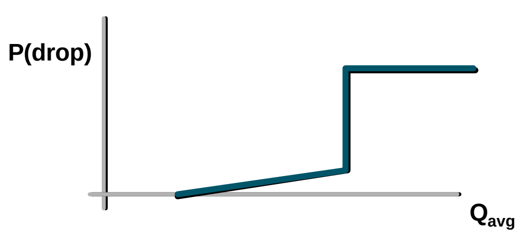
"Colored" (DSCP Written) at Ingress
Classified before Queuing
Potentially Discarded by WRED
Placed in Queues Based on DSCP
Scheduled by CBWFQ (and/or LLQ)

RED Revisited

- **Average queue length is monitored to detect onset of congestion**

  Averaging time constant is O(RTT)

- **As small percentage of packets is dropped, TCP backs off, congestion is averted**



RED Revisited (Cont.)

- **RFC 2309 recommends use of RED to improve network performance**
  - Keep average queues short—buffers are there for the bursts, not to add delay
  - Random drops less likely to force TCP into slow start
- **RFC 2597 recommends RED as basis for Assured Forwarding PHB**
- **RFC 2481 (Experimental) defines Explicit Congestion Notification based on RED**

IPS-430
2891_05_2001_c1 19

Explicit Congestion Notification (ECN)

- **A simple enhancement to RED**
- **Packets are marked rather than dropped (using the other 2 "TOS" bits)**
- **TCP congestion avoidance responds as if drop occurred—other transports may also react appropriately**
- **Congestion avoidance without loss**
- **Host participation required**
  - Packets marked "ECN capable"
  - Receiver conveys marking back to sender at transport layer

IPS-430
2891_05_2001_c1 20

RED: Queue Length Determination

- The queue length in RED is calculated as a weighted moving average

  $L_n = (1 - w) ? L_{n-1} + w ? L_{inst}$

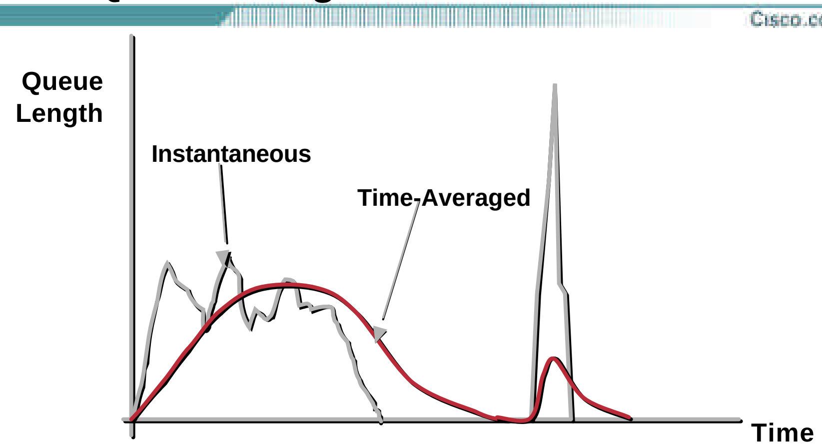
- Time constant reflects typical RTT

  Respond to congestion that can be affected, not to shorter term bursts

- Instantaneous queue measurement alone will drop prematurely, miss congestion
- Extensive analysis and experience of RED assumes moving average approach

RED: Queue Length Determination

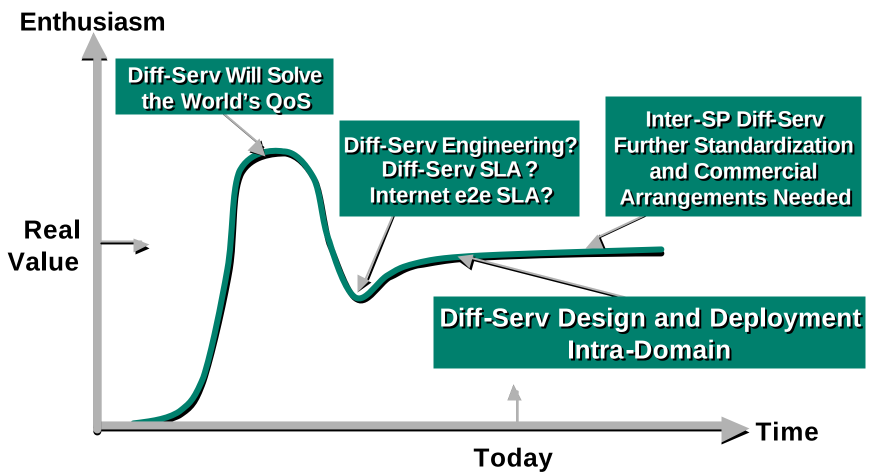
Cisco.com



- Instantaneous reading misses congestion, penalizes bursts

© 2001, Cisco Systems, Inc. All rights reserved.

Diff-Serv Acceptance



Diff-Serv Agenda

- Diff-Serv Recap
- The New "EF" (Expedited Forwarding)
- Worst-Case Bounds and Guarantees
- Services and Per-Domain Behaviors (PDBs)

EF—The Big Picture

Cisco.com

- EF was intended to support services needing guaranteed rate, low jitter
  - (E.g. Virtual Wire, but not only VW)
- RFC 2598 did this by providing guaranteed rate PHB
  - Service rate > arrival rate ? minimal jitter
- draft-charny-…00 pointed out implementation problems for RFC 2598, proposed a fix
- Techno-political chaos ensued…

IPS-430
2891_05_2001_c1 25

RFC 2598 Problems

Cisco.com

- RFC2598 required EF traffic be served at a known output rate R
  - Widely understood as "priority queue or high-weight WRR/WFQ"
- Difficulties arose when dealing with
  - Time interval to measure R
  - Lack of EF packets to serve (maybe due to internal delay)
- Intuitively valid implementations violated spec
  - Appendix was the best part of spec

IPS-430
2891_05_2001_c1 26

Intuition behind New Definition

Cisco.com

- EF packets should ideally be served at rate R (or faster)
- Calculate an “ideal finishing time” for each packet based on “fluid” service at rate R
- Deviation from ideal departure time is bounded by an error term E

  Low E is good

IPS-430
2891_05_2001_c1 27

Ideally, When Should a Packet Leave?

Cisco.com

- Time to serve packet of length L at rate R is L/R

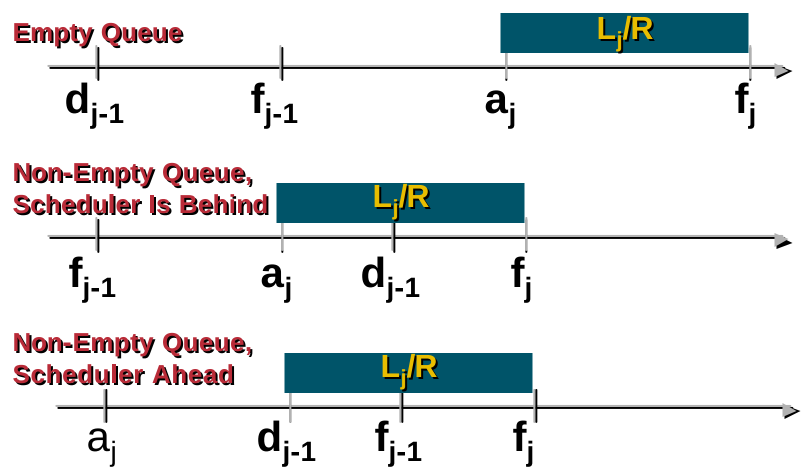
  Service should finish L/R seconds after it starts

- If EF packet arrives at empty EF queue, service should start at once
- If EF packet arrives at non-empty EF queue, service should start right after last packet is served

IPS-430
2891_05_2001_c1 28

The Ideal Finishing Time, f_j

Cisco.com



IPS-430
2891_05_2001_c1 29

The Equations

Cisco.com

- $f_j = \max(a_j, \min(d_j\text{-}1, f_j\text{-}1)) + L_j/R$

  Service is never required to start until packet has arrived

  If scheduler is “late”, service should start after last ideal departure—don’t lose rate

  If scheduler is “early”, service should start after last real departure—don’t penalize better service

  Service should take L/R seconds

- Actual departure lags ideal by E or less:

  d_j ? f_j + E

IPS-430
2891_05_2001_c1 30

“Identity-Aware” Equations

- Original draft bounds delay and jitter only if service discipline is known

  E.g. Can bound per-packet delay and jitter if service was FIFO for EF aggregate

  We didn’t want to mandate FIFO service

  One packet could be delayed for ever in some cases

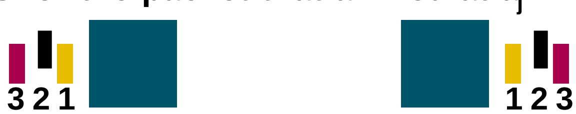
- Solution: Add “packet-identity-aware” version of definition to original one

  Bounds per-packet delay and jitter for any service discipline (given limited input rate and burst)

Aggregate vs. Packet-Identity-Aware

- In aggregate equations, d_j and f_j refer to j^{th} EF departure, which may not be the packet that arrived at a_j

- In new definition, d_j and f_j refer to departure times for the packet that arrived at a_j



Identity-Aware Definition

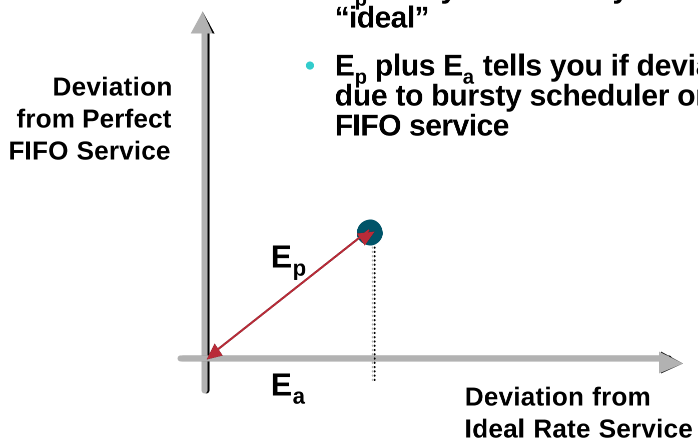
- Looks exactly like old definition, but E becomes E_p (packet error) and a_j and d_j are for the same packet
- Does not require FIFO service of EF packets, but quantifies deviation from “ideal” FIFO EF behavior
- Provides a per-packet delay bound: D_p ? B/R + E_p if traffic offered to interface conforms to (R,B) token bucket

Why Keep the Aggregate Version?

- Closer to RFC 2598
  - Output rate-based definition
- E_p hides many sins
  - E.g. A bursty scheduler and a smooth one with limited misordering look the same
- E_p penalizes certain “reasonable” behaviors
  - E.g. Devices with many ports have large E_p
  - E.g. Per-input WFQ, per “rate-class” queuing

Why E_p and E_a Are Both Needed

- E_p tells you how far you are from "ideal"
- E_p plus E_a tells you if deviation is due to bursty scheduler or non-FIFO service



The New EF—The Bottom Line

- Like RFC 2598, new definition can be implemented by a dedicated queue served at high priority/known rate
  - Without the bugs of 2598
- Two sets of equations allow rigorous conformance testing and quantification of implementation
  - Implementations that come closest to ideal "fluid" service at rate R get best scores
  - E.g. PQ > CBWFQ > WRR

Diff-Serv Agenda

Cisco.com

- Diff-Serv Recap
- The New “EF” (Expedited Forwarding)
- Worst-Case Bounds and Guarantees
- Services and Per-Domain Behaviors (PDBs)

IPS-430
2891_05_2001_c1 37

Disclaimers

Cisco.com

- This section might cause you to become discouraged about Diff-Serv
- This is about worst-case—reality is unlikely to be this bad

  Motivated by desired to understand hard guarantees and to dispel hype
- EF and LLQ remain the recommended approach for low latency
- All the following arguments apply equally well to ATM, FR and IP

IPS-430
2891_05_2001_c1 38

Can Diff-Serv Provide Guarantees?

Cisco.com

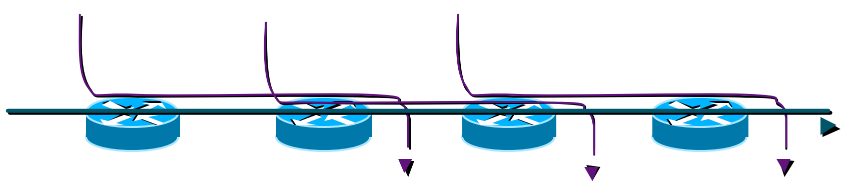
- Confusion around the "virtual wire" (formerly virtual leased line)
- Need some way to bound the offered load on each link

  E.g. ingress policing + topology/traffic matrix knowledge, or admission control
- Even so, "bad" things can happen…
- For analysis, consider a network with only EF traffic in it

IPS-430
2891_05_2001_c1 39

Worst-Case Example

Cisco.com



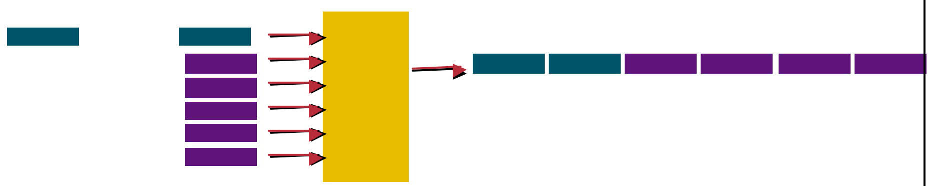
- Burst accumulation is the problem

  Low average utilization but spikes of high utilization
- Smooth "green" flow of rate << line rate intersects many "purple" flows of much lower rate

  Bursts of green result, as we'll see…

IPS-430
2891_05_2001_c1 40

Worst-Case Behavior (1)

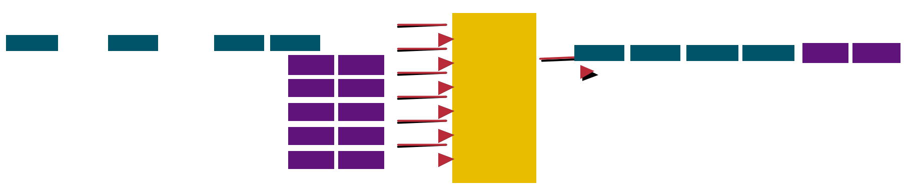


- Green packets arrive well-spaced at rate << line rate
- Purple packets arrive at much lower rate, but synchronized and before first green
  - Output is nowhere near to oversubscribed
- Result: burst of green packets at output

  can make this arbitrarily bad with enough inputs

IPS-430
2891_05_2001_c1 41

Worst-Case Behavior (2)



- Feed outputs from switches like on last page to other switches
- Now bursts of purple packets arrive just before bursts of green
- Result: bursts of green packets get larger
  - Can make this arbitrarily bad with enough hops in the network

IPS-430
2891_05_2001_c1 42

Worst-Case Behavior (3)

- **Suspose**

  We build up a big burst of back-to-back packets on an OC-48

  That reaches a deaggregation router with OC-3 links toward customers

- **Delay experienced by packet at end of burst is almost the time to transmit entire burst at OC-3**

  Effect of lower speed interfaces left to the reader…

Analysis of Worst-Case

- **For general topology, worst case delay bound is known if ? < 1/(h – 1)**

  ? = utilization, h = number of hops

  e.g. ? < 10% for 11 hop network, and delay bound ? ? as ? ? 10%

- **Better bounds for some topologies**

- **Bounds are sensitive to:**

  Utilization

  Burstiness of input streams

  Rates of input streams

Implications

- **Worst-case is improved when**
  - Utilizations are low (duh!)
  - Traffic is well-smoothed on ingress
  - Flows of similar rates grouped in one PHB
    - Recall that burst accumulation followed from a high-rate stream intersecting many low-rate streams
    - Suggests different queues for voice and video, for example

Worst-Case vs. Reality

- **Worst case can theoretically happen—but is it likely?**
- **Relevant experience from ATM CBR**
  - Exact same analysis for ATM with aggregate queuing
  - Input streams typically well smoothed
  - Typical delays and utilizations much better than theoretical worst case (e.g. ? = 50%)
- **Still gaining IP/Diff-Serv experience**
  - Start conservative and ramp up

Diff-Serv Agenda

Cisco.com

- Diff-Serv Recap
- The New “EF” (Expedited Forwarding)
- Worst-Case Bounds and Guarantees
- Services and Per-Domain Behaviors (PDBs)

IPS-430
2891_05_2001_c1 47

Terminology

Cisco.com

- IETF has avoided “services” and “SLAs”

  Notion that each ISP wants to define their own services for differentiation

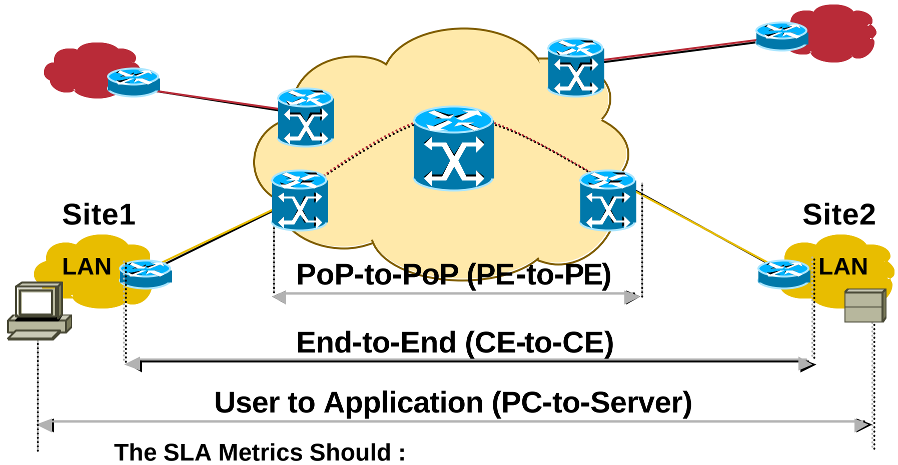
- “Per-Domain Behaviors” (PDBs) are essentially services defined across a single domain

  Virtual wire—the most notorious PDB

- Service level specifications (SLSs) are SLAs without the legal bits

IPS-430
2891_05_2001_c1 48

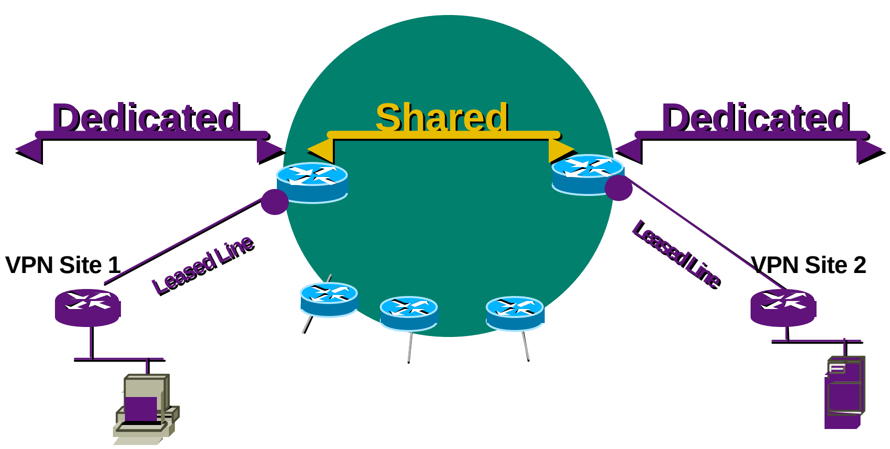
The VPN User's Experience



The SLA Metrics Should :

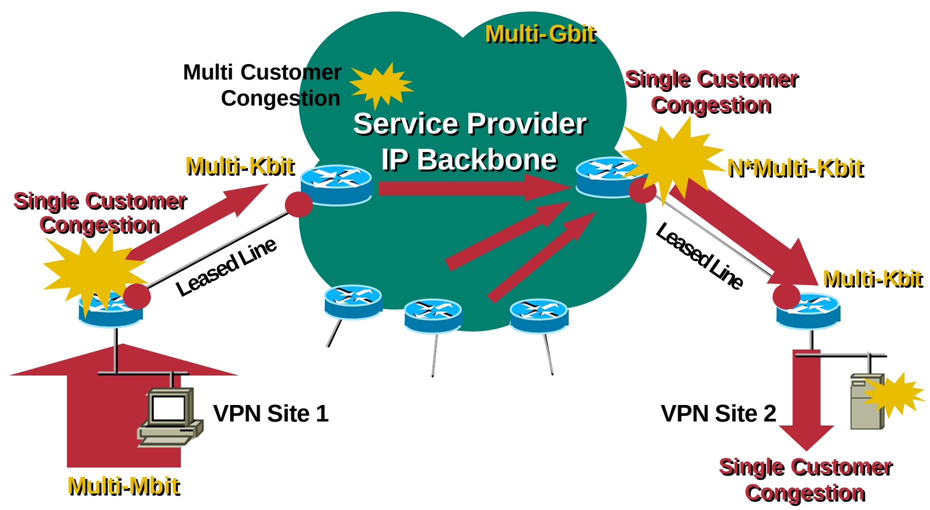
- Show the quality of the shared SP network
- Show the quality of the service provided to the VPN user

Where Is the Customer's Traffic Shared With Other Customers' Traffic ?



Where Can Things Go Wrong? The Potential Congestion Areas

Cisco.com



IPS-430
2891_05_2001_c1 51

Enterprise Customer IP-VPN SLA Requirements

Cisco.com

- Get equivalent or better SLA commitment for IP as for FR, ATM and leased lines services
- Get proactive reporting on the the service performance, including trend analysis and capacity planning
- Motivate service provider to take “problems” seriously (penalties, contract termination…)

IPS-430
2891_05_2001_c1 52

Example SLA Components

Cisco.com

- Bounded delay, 99% delay, or average delay
- Average loss rate
- Service-specific target
  - E.g. transfer time for file of size X
- Availability target
- Statistical rather than worst-case

IPS-430
2891_05_2001_c1 53

How To Deliver SLAs with Diff-Serv

Cisco.com

- EF and AF both allow provider to build logically separate networks
  - EF, AF1, AF2, etc. each run at independently controlled utilization
  - Offered load into a class is controlled by policing
  - Resource allocation per class controlled by queue configuration
- AF also provides means to “shed load”
  - E.g. AF12 is dropped in congestion before AF11

IPS-430
2891_05_2001_c1 54

SLA Delivery with Diff-Serv

Cisco.com

- If you don't like worst-case theory, need to measure
- Run each class at a utilization that meets the SLA target with reasonable safety margin
- Monitor safety margin and violations (e.g. SAA)
- Adapt by reducing load or increasing resources for a class
    - Increase total capacity as the last resort
    - See also traffic engineering

IPS-430
2891_05_2001_c1 55

Cisco.com

RSVP and Integrated Services (Int-Serv)

IPS-430
2891_05_2001_c1 56

RSVP and Int-Serv Agenda

Cisco.com

- **Int-Serv/Diff-Serv Integration**
- **RSVP Aggregation**
- **RSVP Refresh Reduction and Reliability**
- **RSVP Future Directions**

IPS-430
2891_05_2001_c1 57

What Are RSVP and Int-Serv Good For?

Cisco.com

- **Making firm-hard reservations for individual application flows**

  Topology-aware admission control is ideal for voice

- **Main drawback: per flow state in routers**

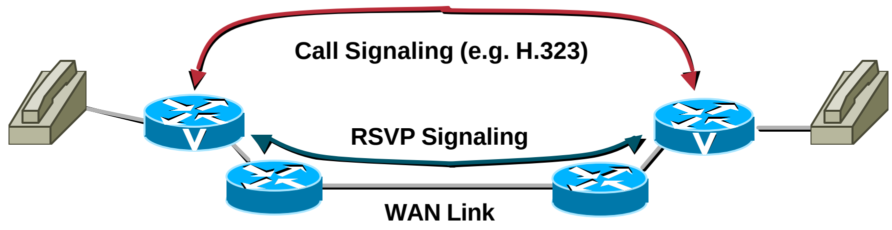
  Scaling properties not attractive to ISPs

- **Valuable today in enterprise nets**

  E.g. Admission control on a WAN link

IPS-430
2891_05_2001_c1 58

RSVP for VoIP



- RSVP support is being built into VOIP gateways today
- Ensure sufficient capacity on WAN links, or notify GW if not
- Avoid admitting the “last straw” call

IPS-430
2891_05_2001_c1 59

Scaling Issues in RSVP

- Classification state: Need to identify each microflow using 5-tuple
- Policer state: Token bucket per microflow
- Scheduling state: Guaranteed service needs a queue per microflow
- Reservation state: Data structures for each reservation

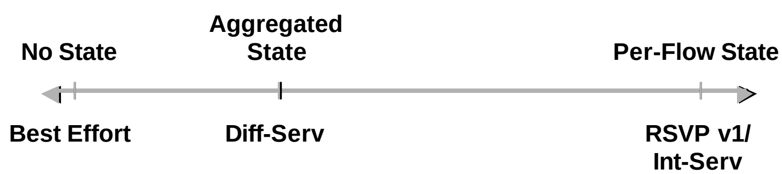
  Both storage and refresh costs

IPS-430
2891_05_2001_c1 60

Differentiated Services

Cisco.com

- Created largely in response to scaling concerns about RSVP



IPS-430
2891_05_2001_c1 61

The Trouble with Diff-Serv

Cisco.com

- As currently formulated, Diff-Serv is strong on simplicity and weak on guarantees
- Virtual wire using EF is quite firm, but how much can be deployed?

  No topology-aware admission control mechanism

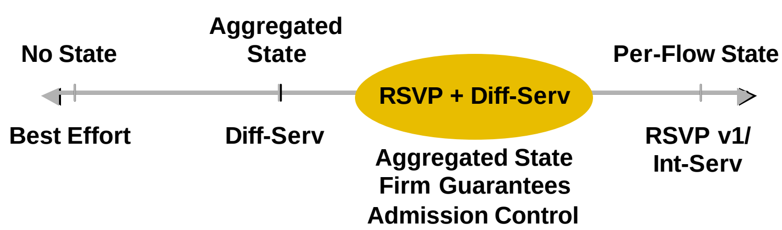
- Example: How do I reject the “last straw” VOIP call that will degrade service of calls in progress?

IPS-430
2891_05_2001_c1 62

RSVP/Diff-Serv Integration

Cisco.com

- The best of both worlds...



IPS-430
2891_05_2001_c1 63

Problem Statement

Cisco.com

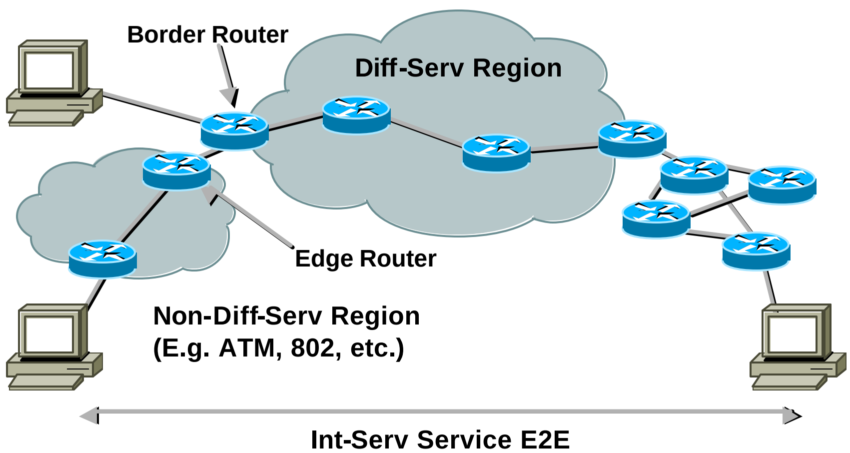
- **Combine Int-Serv end-to-end for firm guarantees with Diff-Serv in core for scalability**

  Given the presence of a Diff-Serv “cloud” in a network that aims to support Int-Serv E2E, how do we meet the QOS goals of applications?

  Analogous to handling ATM, 802, etc. clouds in Int-Serv networks

IPS-430
2891_05_2001_c1 64

Network Model



IPS-430
2891_05_2001_c1 65

Addressing Scaling Issues

- **Per-flow queuing**

  Use controlled load or LLQ

- **Reservation state**

  Aggregate RSVP and refresh reduction

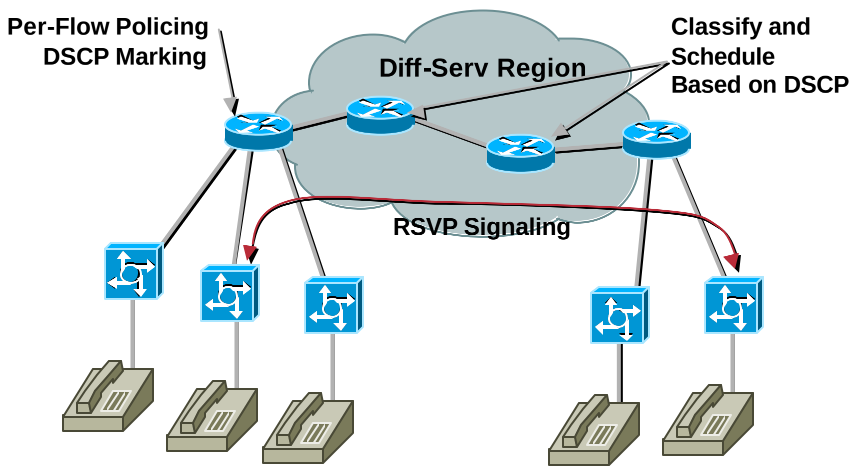
- **Per-flow classification and policing**

  Use Diff-Serv data plane with Int-Serv control plane, i.e. classify with DSCP

IPS-430
2891_05_2001_c1 66

RSVP/Diff-Serv

Cisco.com



IPS-430
2891_05_2001_c1 67

RSVP/Diff-Serv Step 1

Cisco.com

- Routers at edge of DS cloud perform microflow classification and policing, set DSCP
  - May use 2 values for “in” and “out” of contract
  - Guaranteed ? EF, controlled load ? AFx is an option, but it’s a local policy choice
- RSVP is used at every hop for admission control
- DSCP classification and scheduling in the core
- Scaling limit is now just the per-flow reservation state
  - Control plane only
  - > 10k reservations per node already demonstrated

IPS-430
2891_05_2001_c1 68

RSVP and Int-Serv Agenda

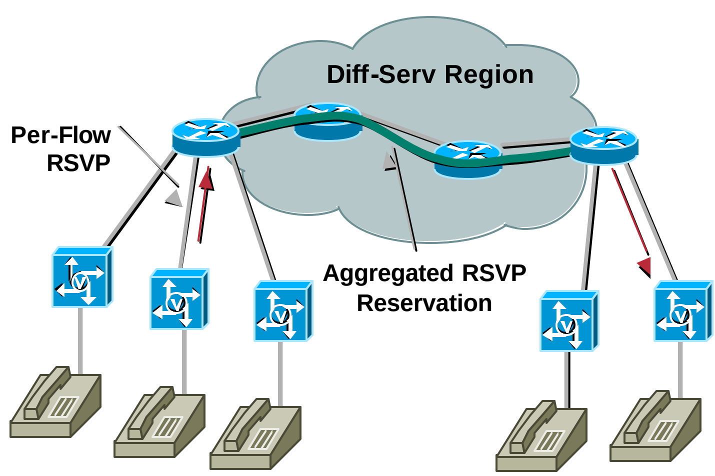
Cisco.com

- Int-Serv/Diff-Serv Integration
- **RSVP Aggregation**
- RSVP Refresh Reduction and Reliability
- RSVP Future Directions

IPS-430
2891_05_2001_c1 69

RSVP/Diff-Serv (2)

Cisco.com



IPS-430
2891_05_2001_c1 70

Presentation_ID.scr

RSVP Aggregation

- Forwarding plane: Still Diff-Serv
- We now make aggregated reservations (“fat pipes”) from ingress to egress
- Microflow RSVP messages are “hyperspaced” across cloud
- Size of aggregate reservation may be dynamically adjusted to cover all microflows
  - Heuristics are possible

IPS-430
2891_05_2001_c1 71

RSVP Aggregation Details (1)

- E2E RSVP messages represent microflows
- Aggregate RSVP messages represent “fat pipes” of many flows
- Aggregation region is created by configuring routers to aggregate and de-aggregate
  - Such routers have “interior” and “exterior” interfaces
  - Aggregation occurs when E2E path goes from interior to exterior

IPS-430
2891_05_2001_c1 72

RSVP Aggregation Details (2)

Cisco.com

- Aggregation router swaps IP protocol number from “RSVP” to “RSVP-E2E-IGNORE”

  Ignored by core routers

  Intercepted and restored to “RSVP” by de-aggregation (egress) router

- Egress router send PathErr back to ingress

  End-points for new aggregate reservation are thus auto-discovered

IPS-430
2891_05_2001_c1 73

RSVP Aggregation Details (3)

Cisco.com

- Ingress and egress routers build an aggregate reservation

  Like normal reservation, but new session type identifies the DSCP

  All flows sharing common ingress, egress and DSCP belong to one aggregate session

  Size of aggregate is determined by summation of E2E paths and resvs—heuristics may be used to reduce churn

IPS-430
2891_05_2001_c1 74

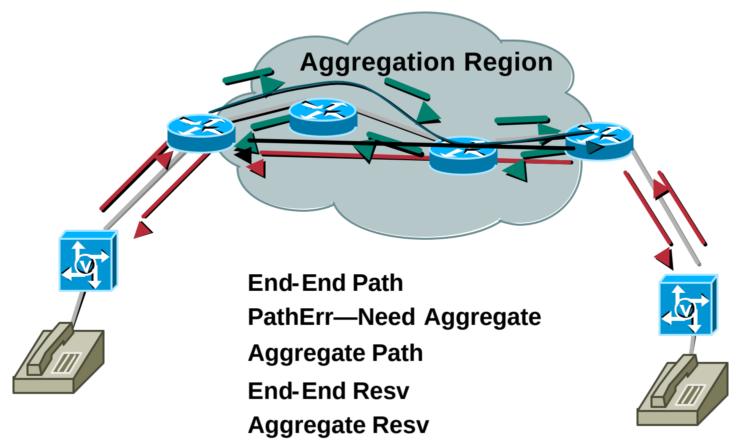
Path Consistency

Cisco.com

- **Aggregate reservation is built by sending path from ingress to egress routers**
- **Path carries destination address of egress; data carries DA of true endpoint**
- **In most cases, path and data will go the same way, but it can't be guaranteed**
  - E.g. Equal cost load balancing
- **Best solution is to force the data to follow the Path, e.g. by tunneling**
  - MPLS is one option here

RSVP Aggregation

Cisco.com



Status

Cisco.com

- RFC has (finally) passed IESG review
- In the RFC editor’s queue
- Partial implementation in IOS today

  More on this shortly…

IPS-430
2891_05_2001_c1 77

RSVP and Int-Serv Agenda

Cisco.com

- Int-Serv/Diff-Serv Integration
- RSVP Aggregation
- RSVP Refresh Reduction and Reliability
- RSVP Future Directions

IPS-430
2891_05_2001_c1 78

Refresh Reduction Issues (1)

Cisco.com

- **Concern about the cost of refreshing large numbers of RSVP reservations**
- **Increasing timer values would reduce reliability**
- **Even at default values, lack of message reliability could be a problem**

  One lost message could delay reservation establishment by minutes

IPS-430
2891_05_2001_c1 79

Refresh Reduction Issues (2)

Cisco.com

- **Debate centered around steady-state load**
- **Primary issue in scaling signaling is dealing with the impulse load under transient conditions**
- **Nothing to do with refresh/soft state—best handled by sound implementation**

  E.g. Message pacing, efficient processing of common cases, etc.

IPS-430
2891_05_2001_c1 80

RSVP “Refresh Reduction”

Cisco.com

Provides:

- Message reliability
- Faster state update
- Resynchronization
- Reduced message load

IPS-430
2891_05_2001_c1 81

Refresh Mechanisms

Cisco.com

- Designed to retain multicast support
  - Running over TCP would lose this
- Message ID
  - Message ID Ack
  - Message ID Nack
- Summary refresh
- Piggybacking

IPS-430
2891_05_2001_c1 82

Trigger and Refresh Messages

Cisco.com

- Most RSVP refreshes are exact copies of previously sent messages
- A message which differs from the previous message is called a trigger message
- Trigger messages are always sent with a new Message_ID
- Refresh messages are sent with the old Message_ID (if they're sent at all)

IPS-430
2891_05_2001_c1 83

Reliability and Responsiveness

Cisco.com

- Message_ID object has an Ack request flag
- Receiver must Ack if this is present
- Prior to receiving an Ack, the sender may use a shorter refresh timer to ensure timely delivery

IPS-430
2891_05_2001_c1 84

Summary Refresh

Cisco.com

- Once a message has been sent with a Message_ID and acked, it can be refreshed simply by sending the Message_ID
- The summary refresh message carries a list of message IDs to be refreshed

IPS-430
2891_05_2001_c1 85

Receiver Side State Sync

Cisco.com

- Senders periodically refresh all messages within time-out period of those messages
- If receiver finds an unknown Message_ID, send Nack for that message
- Sender response to Nack is to send the full message

IPS-430
2891_05_2001_c1 86

RSVP and Int-Serv Agenda

Cisco.com

- Int-Serv/Diff-Serv Integration
- RSVP Aggregation
- RSVP Refresh Reduction and Reliability
- RSVP Future Directions

IPS-430
2891_05_2001_c1 87

RSVP Futures

Cisco.com

- IETF is considering future directions in signaling
- RSVP possibilities
  - Enhanced resource sharing
  - Two-phase operation

IPS-430
2891_05_2001_c1 88

Enhanced Resource Sharing

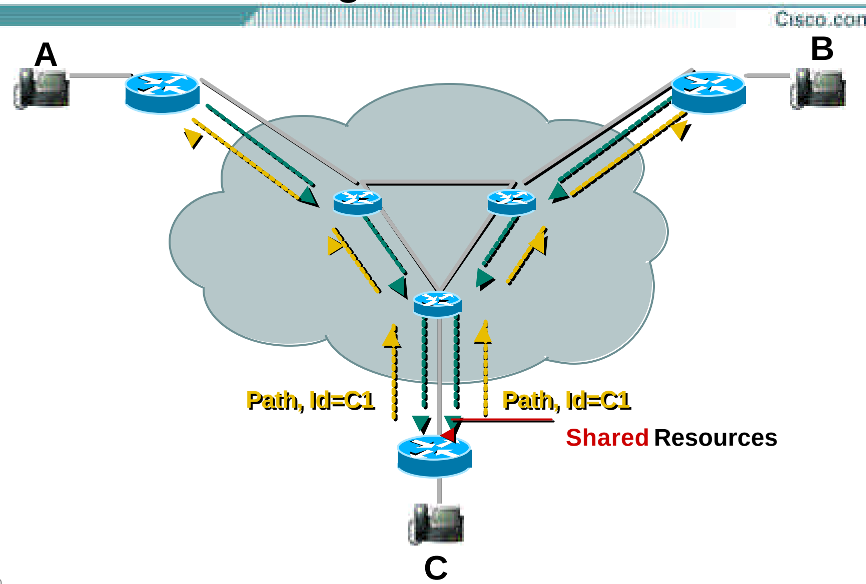
- **RSVP allows sharing among multiple senders, but what about multiple receivers?**

  E.g. In call waiting, why book two sets of resources when I can only talk to one person

  Should be able to reserve one set of resources, shared among calls, under control of application

  Basic approach: A "session ID" to identify sessions (calls) that may be shared

Resource Sharing



Two-Phase Operation

Cisco.com

- **RSVP reserves resources and makes them available at once**
- **In telephony, reserve resources before ringing, but make them available only when phone is answered (and billing starts)**
- **Suggests a 2-phase approach**

  Second phase “commits” resources and atomically starts billing

Cisco.com

The Role of MPLS in QoS

What Role for MPLS?

Cisco.com

- “MPLS brings the QoS of ATM to IP”
- “MPLS is just a traffic engineering tool”
- “MPLS is pure evil”

IPS-430
2891_05_2001_c1 93

“MPLS Brings the QoS of ATM to IP”

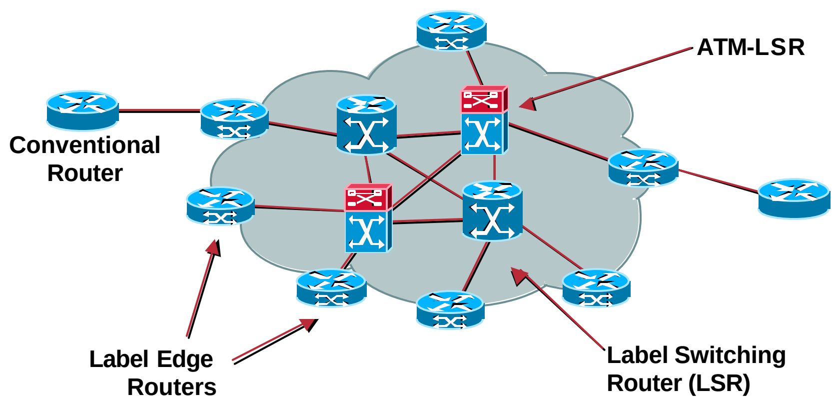
Cisco.com

- ATM has few mechanisms that aren’t in either Diff-Serv or Int-Serv
  - Admission control, policing, class-based queuing, per-flow queuing,…
  - The lack of labels (or VCs) isn’t the obstacle to QoS in MPLS
- In this regard, MPLS QoS (and ATM QoS) has been oversold

IPS-430
2891_05_2001_c1 94

Network Architecture

Cisco.com



End-to-End Service Is (Usually) IP

? IP Quality of Service Is What MPLS Must Support

IPS-430
2891_05_2001_c1 95

MPLS QoS or IP QoS?

Cisco.com

– MPLS is (usually) not end-to-end

– MPLS does not change the service interface

+ MPLS enables constraint-based routing

+ LSPs assist in QoS scalability

Can associate resources (e.g. a queue) with an LSP carrying aggregated traffic

IPS-430
2891_05_2001_c1 96

MPLS and IP QoS

- First order of business for MPLS: Support the complete IP QoS model
  - Diff-Serv—draft-ietf-mpls-diff-ext-09.txt
  - Int-Serv—RSVP extensions allow labels to be bound to reservations
  - Initial goal: Neither more nor less than IP QoS

IPS-430
2891_05_2001_c1 97

RSVP/Diff-Serv Integration with MPLS

- Diff-Serv packet marking provides the key to increasing RSVP scalability
- This could be done without MPLS, but MPLS adds
  - Constraint-based routing of reservations
  - Consistent routing of reservation messages and data
  - Variable granularity of reservations—LSP is the unit of reservation, not microflow or DSCP

IPS-430
2891_05_2001_c1 98

Constraint-Based Routing

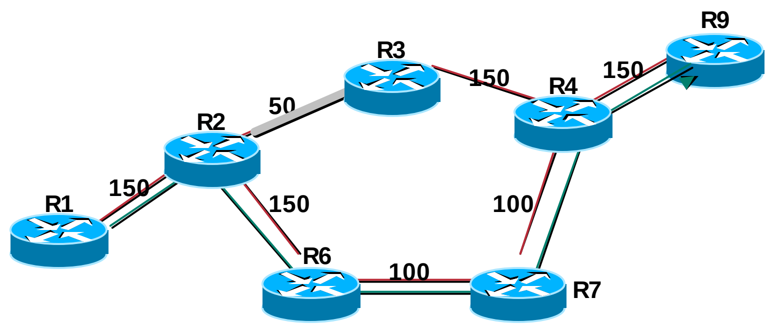
- **IP QoS has typically assumed complete separation of routing and QoS**

  Routing determines the path, QoS determines resource allocation on the path

- **What about picking a path with appropriate resources?**

  Constraint-based routing: Picking a path that meets certain constraints (e.g. sufficient BW, low delay)

Constraint-Based Routing Example



Trying to Find a Path from R1 to R9 with Bandwidth 75 Mbps
R2-R3 Link Violates Constraint (BW ? 75) So Delete It
Pick Shortest Path on Remaining Topology

Is Current IP Routing Sufficient?

Cisco.com

- Conventional IP routing distributes consistent view of network to all nodes in an area
- In constraint-based routing, packets from different sources may need to be forwarded according to different constraints

IPS-430
2891_05_2001_c1 101

Is Current IP Routing Sufficient? (Cont.)

Cisco.com

- Conventional IP routing uses pure destination-based forwarding
- In constraint-based routing, packets from different sources may need to be forwarded according to different constraints
  - Need some “source routing” capability
  - IP source route option has limitations

IPS-430
2891_05_2001_c1 102

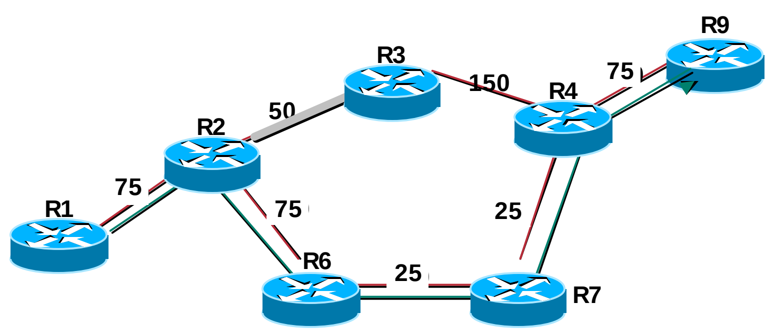
Role of MPLS

Cisco.com

- An MPLS LSP can be explicitly routed along a path that meets the constraints
  - Using explicit route object + label object in RSVP (or CR-LDP)
- Resources may be allocated at time of LSP establishment

Constraint-Based Routing Example

Cisco.com



Trying to Find a Path from R1 to R9 with Bandwidth 75 Mbps
R2-R3 Link Violates Constraint (BW ? 75) So Delete It
Pick Shortest Path on Remaining Topology
Update Available Capacities When Path Is Established

MPLS Benefits and Limitations

- **Current implementation is limited to single area, link-state protocol**
  - Multi-area work underway
  - Inter-domain and non-link-state are challenging
- **A good match to RSVP aggregation**
  - Aggregate reservation limited to a domain
  - MPLS ensures PATH and data follow same route

Status

- **MPLS "guaranteed bandwidth" and "Diff-Serv-Aware Traffic Engineering" (DS-TE) in 12.0(11)ST and 12.2(2)T**
- **IETF efforts:**
  - RSVP extensions for MPLS past last call
  - Extensions to RSVP and IGPs for DS-TE in progress

Concluding Remarks

Cisco.com

IPS-430
2891_05_2001_c1 107

Conclusions

Cisco.com

- IP QoS is not an oxymoron
- There are many useful IP QoS technologies to deploy today

  Diff-Serv: Scalability for large networks

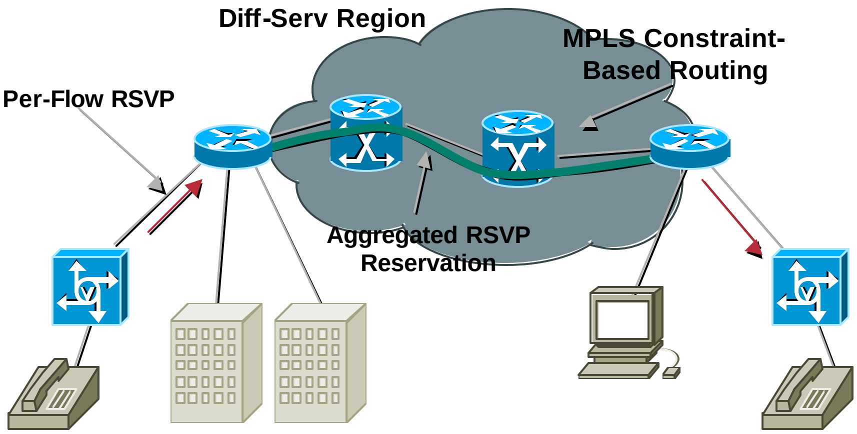
  RSVP and Int-Serv: For stronger guarantees

  MPLS: For constraint-based routing in backbones (and for non-QoS applications)

IPS-430
2891_05_2001_c1 108

Each Technology Has a Role

Cisco.com



IPS-430
2891_05_2001_c1 109

Forthcoming Capabilities

Cisco.com

- Diff-Serv kinks to be worked out and deployment experience gained
- More analysis of guarantees
- RSVP becoming more scalable and feature-rich
- RSVP/Int-Serv and Diff-Serv will be more closely coupled

  Scalability with firm guarantees and admission control

IPS-430
2891_05_2001_c1 110

Cisco.com

Advanced Concepts and Developments in Quality of Service

Session IPS-430

IPS-430
2891_05_2001_c1 111

Cisco.com

Please Complete Your Evaluation Form

Session IPS-430

IPS-430
2891_05_2001_c1 112

CISCO SYSTEMS

EMPOWERING THE INTERNET GENERATION℠

IPS-430
2891_05_2001_c1 113